



TITLE:

<研究論文>学力調査の時代を読み解く評価リテラシー

AUTHOR(S):

石井, 英真

CITATION:

石井, 英真. <研究論文>学力調査の時代を読み解く評価リテラシー. 教育方法の探究 2019, 22: 1-10

ISSUE DATE:

2019-03-25

URL:

<https://doi.org/10.14989/241657>

RIGHT:

許諾条件により本文は2020-03-26に公開

学力調査の時代を読み解く評価リテラシー

石井 英真

1. はじめに

2000 年前後の学力低下論争以降、国内外の学力調査が、間接的にあるいはより直接的に教育政策や教育実践に影響を及ぼしてきた。本稿では、日本における学力向上政策を中心に、その中で学力調査の位置づけ、および、その背景と論点を素描するとともに、学力調査を飼いつづけるのに必要な知（評価リテラシー）の重要性とその中身について論じる。

2. 日本における学力向上政策と学力調査の展開

（1）学力低下論争と全国学力テストの始まり

1999 年、大学生の学力問題から始まった学力低下論争は、「ゆとり教育」を標榜する従来の教育課程政策の是非を問い、さらに、OECD の PISA2003 で、読解力の順位が 8 位から 14 位に下がったことは、教育界に「PISA ショック」をもたらした。これにより、「ゆとり教育」から「学力向上」への方針転換は決定的なものとなった。2008 年には小・中学校の学習指導要領が改訂され、知識・技能の習得と、PISA 型学力を意識した「活用する力」を車の両輪として重視することが示された（「確かな学力」観）。

学力向上をめざす文部科学省は、2007 年 4 月に全国学力・学習状況調査（全国学力テスト）を実施するようになった。全国学力テストは、小学校 6 年生、中学校 3 年生を対象にした悉皆調査（2010 年度、2012 年度は抽出調査及び希望利用方式、2011 年は東日本大震災の影響を考慮し調査としては実施せず）であり、国語と算数・数学について調査される（2012 年度、2015 年度、2018 年度調査では理科も実施、2019 年度は中学校で英語も実施予定）。調査問題は「知識」を問う A 問題と「活用」を問う B 問題に大別され、B 問題は後述する PISA を意識したものになっている（2019 年度調

査からは、A 問題、B 問題の区分を廃止し、知識と活用を一体的に問う問題構成に改められる）。教科の学力実態の調査に加え、児童・生徒の生活習慣や学校環境を問う質問紙調査も実施される。

（2）全国学力テストの結果の活用をめぐる状況

こうした学力テストの実施は、教育制度の構造改革とも密接に関わっている。規制緩和により教育現場の創意工夫を促す一方、学力調査（学力テスト）を実施するなどして、学校教育の質保証のシステム（PDCA サイクル）を確立することで、教育の成果への統制を強めるわけである。

調査結果に関しては文科省が集計し、都道府県別の平均正答率が公表される。市町村別や学校別の結果も集計される。学校間競争やテスト準備教育を煽るものとなる危惧から、データの提供を受けた各教育委員会が個別の市町村名や学校名を公表することはこれまで禁じられてきた。しかし、2014 年度調査から、結果を分析して改善策を示すことを条件に解禁した（平均正答率を一覧表にすることや順位付けは認めないが、ルール違反への罰則はない）。都道府県教育委員会も市町村教委の同意があれば、市町村別や学校別の成績を公表できることになった。

全国学力テスト導入当初には、愛知県犬山市が自治体レベルで不参加を貫いていたが¹、2009 年度以降、一部の私立学校を除きほぼすべての小・中学校が参加するようになった。また、全国学力テスト導入の前後から、都道府県や市区町村のレベルで、各自治体固有のテストも実施されており、子どもたちは学力テストを頻繁に受けるようになっている²。さらに、実際に学校別成績を公表しているのは大阪市、松江市、武雄市など少数派ではあるが、都道府県知事や市区町村長を

中心に、情報公開や説明責任のために結果を公開すべきとの主張もなされ続けている。

3. 学力調査の時代の背景と論争点

(1) 全国学力テストに対する危機

かつて文部省によって、1956～66 年にかけて全国一斉学力調査が行われたが、1961～64 年に中学 2・3 年生に対して実施された学力調査が悉皆調査であったことから、「学力テスト反対運動」（学力テストが、児童生徒、教師、学校の管理や競争の道具となることへの批判）が起こり終了した。同様に、昨今の日本の学力向上政策に対しても、「新自由主義」教育改革、「学力テスト体制」などという言葉で、批判がなされている³。すなわち、現代日本の教育改革は、テスト学力とテスト準備教育を重視する傾向をもたらし、教育における競争主義と成果主義（結果至上主義）を強化するというのである。

上記のような日本における教育改革の展開、すなわち、学力向上が課題となり、それを実現する方策として学力テストが用いられている点は、英米を中心に世界的に展開を見せている「スタンダードに基づく教育改革（standards-based reform）」を想起させる。たとえば米国では、「どの子ども置き去りにしない（No Child Left Behind : NCLB）」法以降、州統一の学力テストの結果を、学区や学校への予算配分、教職員の処遇、保護者による学校選択と結び付ける事態が全米規模で展開している。こうして、学力テストが子どもや学校にとって大きな利害のからむ「ハイ・ステイクス」なテストとしての性格を強めるに伴って、テスト結果を上げることが至上命題となり、学校間競争が激化し、授業がテスト準備に矮小化されたり、教育困難校や低学力の子どもたちの切り捨てが進行したりしている⁴。

(2) 世界的な学力テスト政策の展開の背景

現代日本の学力テスト政策は、世界的に展開している「スタンダードに基づく教育改革」の一つの形と見ることができる。「スタンダードに基づく教育改革」の展開は、グローバル化の進展と福祉国家の縮小を背景とした、公教育の新たなガバナンスの形の模索という、先進諸国に共通する課題と関連している。社会の成熟段階において、国家が福祉国家のコストを担うことが

財政的に困難となっており、しかも変化が激しく人々のニーズも多様化する現代社会の中では、各地域の固有の状況に応じる柔軟性が公共部門にも求められるようになる。その結果、これまで国家が画一的な形で担ってきた公共領域を、さまざまなエージェントが分担して担っていくことが求められるようになっており、その際、市場に委ねるか（新自由主義的方向性）市民社会に委ねるか（社会民主主義的方向性）という点が、そして、局所的で多様な取り組みの質をどう保証するかという点が問われるようになってきているのである。

また、低成長時代、人口減少社会においては、国や地方の財政状況がきわめて厳しくなる。その中で、教育分野への財政支出を増やすにしても削減するにしても、財政支出に見合った実践や政策の効果を誰もが納得できる根拠に基づいて説明することが求められるようになっている。たとえば、文科省と財務省との関係、教育関係者と一般市民との関係など、教育を専門としない人々をどう納得させるかが問われている。そうした状況下で、説明責任の要求に応える一つの方法として、「スタンダードに基づく教育改革」は「エビデンスに基づく教育（evidence-based education）」を強調する傾向にある。

しかも、PISA をはじめとした OECD の「教育インディケーター事業（Indicators of Education Systems）」が、政策評価に止まらず、各国の教育システムを同一基準により比較し、教育改革を方向付けるようになる中で、「エビデンスに基づく教育」との結びつきを深めながら、「スタンダードに基づく教育改革」という改革手法のパッケージ化・標準化が加速している⁵。

(3) 現代日本の教育改革における学力テストの位置

今世界各国で、「21 世紀型学習」「総合型学力」「コンピテンシー」など、21 世紀の社会で求められる包括的かつ汎用的な学力観が打ち出され、それを系統立てて育てるべく、カリキュラム改革と評価改革がセットで進行している。一方で、学力テストの活用の仕方については、さまざまな形がある。アメリカのように、テスト結果による制裁が規定されていて、結果に基づいて教師や学校が厳しく値踏みされる国もあれば、オーストラリアのように、結果は公表しても支援的な介入を重視している国もあれば、イタリアのように、悉

皆調査を行いつつ全国的な学力水準を確認するためにサンプルを抽出・分析し、あとは学校の自己評価に役立てさせる国もあれば、スウェーデンのように、学校間の評価の質の調整や学力観の共有に役立て、教師の自主的な活用を尊重している国もある。学力テストを活用する方法は、選抜か説明責任か教育改善かという諸目的のどれを重視するかによって、様々な形が考えられるのである⁶。

では、日本の教育改革において学力テストはどのような役割を担っているのか。それは教育現場に何をもたらしているのか。まずは、日本において学力テストがどのような位置づけにあるのかを確認しておこう。志水宏吉らは、全国学力・学習状況調査の教育現場への影響について検討し、下記のような状況を明らかにしている⁷。

全国学力テストが実施され、都道府県別の結果が公表されたことによって、日本の「学力地図」が明らかになり、都道府県レベルでの競争は過熱している。そして、各都道府県で学力水準の向上が重視されるようになり、全国学力テストに対応する学力向上策が実行されるようになった。

ただし、こうした全国学力テストの都道府県の教育政策への影響には濃淡が見られる。全国学力テストのインパクトが最も大きいのが沖縄、高知、大阪などの「下位県」であり、秋田、福井、香川などの「上位県」がそれに次ぎ、もっとも影響が少なかったのが、神奈川、兵庫、宮崎などの「中位県」である。また、「下位県」では、府県の教育委員会や首長の存在感が増しているという。テストによる教育の国家統制の強化が危惧される一方で、地方独自の教育課題や教育の歴史を反映して、教育の政策が地域ごとに多様化している事実を志水らは指摘している。

全国学力テスト実施前から、地方自治体レベルでも多くの悉皆の学力テストがなされてきた。しかし、一部の自治体を除けば、保護者や地域住民などへの、市町村別、学校別の学力テストの結果の公表は限定的である。また、学力テストの結果については、間接的な影響は大きいにしても、それを学校選択、あるいは、個々の学校や教職員への処遇と直接的に結び付けている自治体はほとんどない。結果の公表が認められる中、競争主義や成果主義を強化する方向で学力テストの利

用が広まることには注意が必要である。ただ、英米と比べれば、現時点で日本の学力テストは必ずしもハイ・ステイクスなものとはなっていない。

むしろ日本において、全国学力テストは、実質的には、学習指導要領の学力観の趣旨を直接的に伝達・徹底する機能を果たしてきた。その結果、テスト準備教育といっても、知識・技能の習得のみならず、知識・技能を活用する力にも強調点が置かれている。それは、考える活動を重視する授業を促す一方で、思考力・判断力・表現力の育成が、「活用」という型をなぞる学習に矮小化され、形式化される危険性もはらんでいる⁸。また、学力向上の手立てが授業改善に一元化されることで、学校行事等の教科外活動の縮小が進んでいる点にも注意が必要である。

さらに、実施された学力テストについて、自治体や学校は、結果の分析とそれに基づく改善というアクションを取ることが求められる。こうして、「PDCA サイクル」という言葉と形式は、学校経営はもちろん、授業づくりのレベルにまで浸透している。以上のように、日本において学力テストは、英米ほどには、教育の結果を厳格に管理・統制するものとはなっていない一方で、教育の目標・内容、および、教育実践の方法を、特定のあり方へとより直接的に枠づけるものとして機能している⁹。

4. 学力調査の時代に必要な評価リテラシー

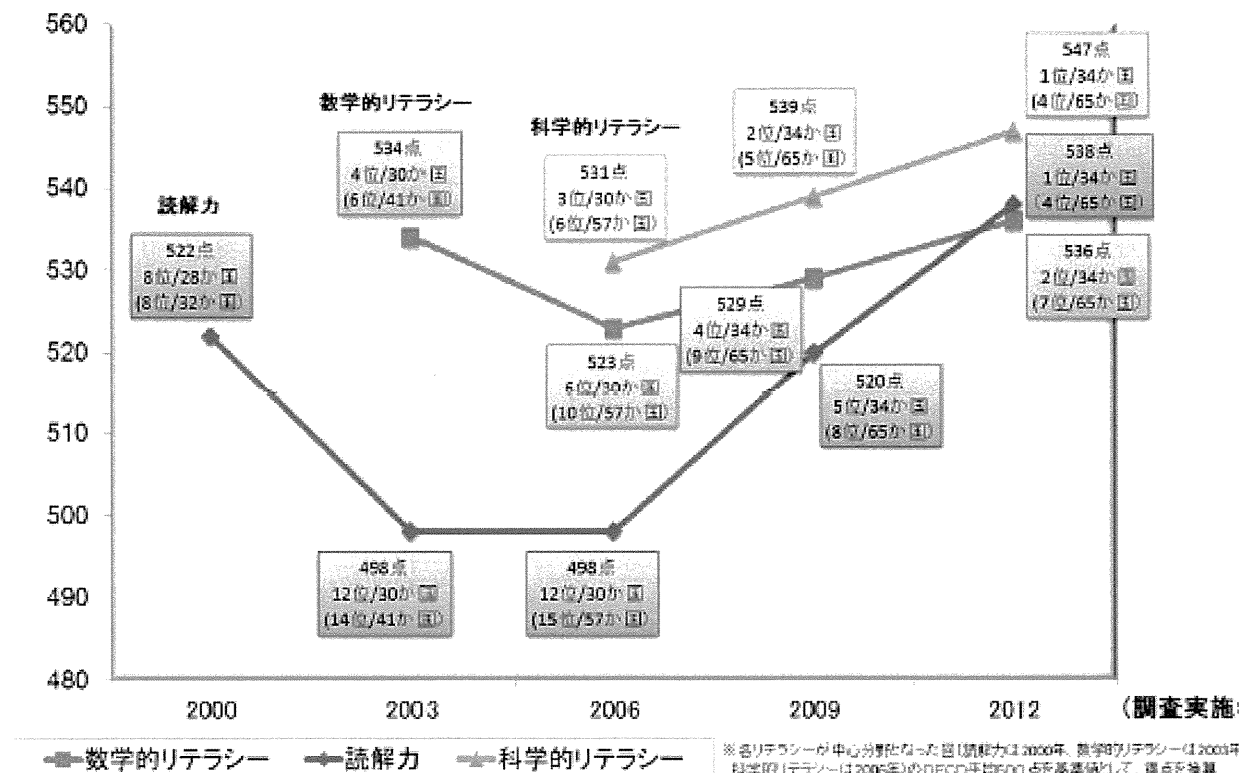
（１）学力調査の時代に求められるもの

学力テストや「スタンダードに基づく教育改革」を批判しそれに対するオルタナティブを探ろうとする者の多くは、「スタンダードに基づく教育改革」という発想や学力テスト自体を、さらには、目的・手段関係や、目標を明確化しそれに基づいて評価すること自体を否定しがちである。しかし、成果を求めるからといって、手続きの効率のみを追求するとは限らない。また、教育実践が本質的に不確実な性格を持つにしても、何らかの見通しなしに教師は教育実践を展開できない。何より、子どもたちの学習権を保障すべく、目の前の子どもたちに何が必要なのかを問い、学習の結果に責任を持つとすることは、教師たちの自発的な応答責任に由来する。

しかも、現在の問題状況は、目標と評価のあり方に

資料1. 平均得点及び順位の推移

(平均得点)



※順位は OECD 加盟国中 (カッコ内は全参加国・地域中の順位)

※数学的リテラシー、科学的リテラシーは経年比較可能な調査回以降の結果を掲載

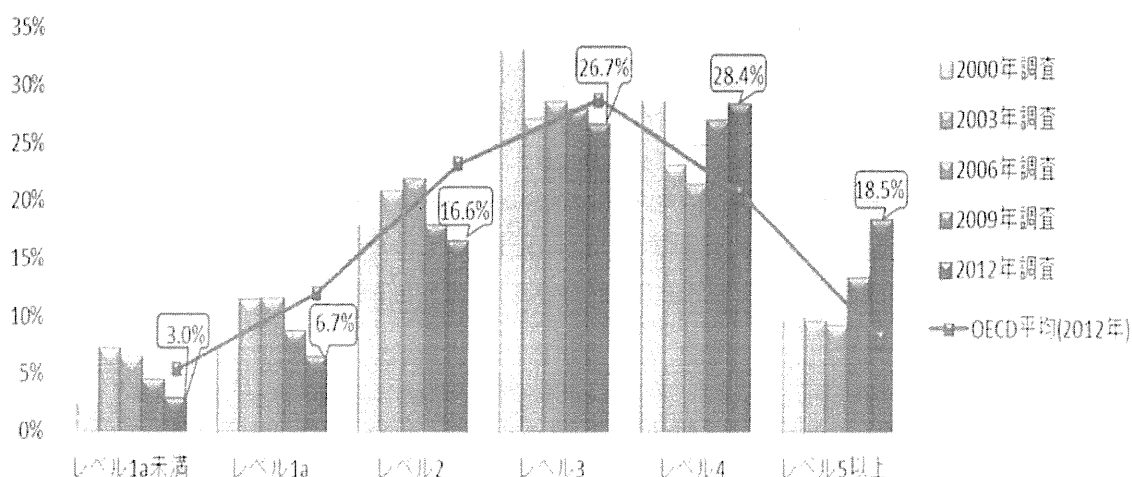
なお、数学的リテラシーについては、2000 年度調査では 557 点で 1 位。科学的リテラシーについては、2000 年度調査では 550 点 2 位、2003 年度調査では 548 点で 2 位であった (いずれも順位は全参加国・地域中)。

について議論・決定する権限の縮小による、教師の専門職性・自律性の危機と密接に関連している。教育を画一化・規格化する「標準化 (standardization)」の作業は、「スタンダード (standards)」(共通教育目標) を明らかにしていく営みとはイコールではない。目指す人間像・社会像を問う、教育的価値に関する公共的議論のフィールドとして、教育目的・目標設定と評価のあり方に関する議論を捉えることがまずは必要である。

学力調査の時代において必要なのは、目標と評価に関する議論を避けるのではなく、調査結果を冷静に読み解き、教育における公正を追求する道具として活用するための、専門的知見の蓄積と教育関係者の「評価リテラシー」の育成である¹⁰。「学力テスト・

評価アレルギー」に陥ることなく、測定論や調査論の蓄積にも学びながら、教育の条理に即した評価研究を進めることが重要なのである。1960 年代の学力テスト反対運動は、学力テストや評価そのものを否定する傾向を生み出し、それらを教育学的議論の外側に置くことになった。学力評価、教員評価、学校評価などの一連の評価政策が教育実践にもたらしている混乱は、教育関係者の「評価リテラシー」の未成熟にも一因があるのではないだろうか。この点に関わって、木村拓哉は、日本において学力調査を担う「テスト専門家」が、「教科の専門家」と「教育心理学者 (教育評価論者)」を軸に構成されており、科学的に厳密な方法論に基づいた調査を行うのに必要な専門家 (「教育測定 (テスト理論) の専門家」

資料 2. 日本の習熟度レベル別の生徒の割合（経年変化）（読解力）



「サンプリングの専門家」「情報処理の専門家」）を欠いた状況であることを指摘している¹¹。

以下、学力調査の結果の解釈と活用に関して、それを教育の効果的な改善につなげていくための知見（教育学的観点からの評価リテラシー）を提起する¹²。具体的には、学力調査の結果の解釈に関わる知（学力調査の結果を読み解く 4 つの視点と背景要因への配慮）、学力調査の結果の活用に関わる知（実践的判断におけるエビデンスとしての調査結果の位置づけ）、学力調査の再設計に関わる知（学力調査の性格規定と設計方針）の三つの観点で見ていくことにしよう。

（2）学力調査の結果を読み解く 4 つの視点

学力調査の結果が公表されると、得点や順位の高低にのみ注目が集まりがちである。しかし、学力調査の結果から教育活動の改善への示唆を得るには、より多面的・複合的に結果を読み解く必要がある。すなわち、学力水準（当該集団の平均値が高いか低いか）のみならず、学力構造（どのような学力の質や中身に課題があるのか）、学力格差（子どもたちの学力がどのように分散しているか）、さらには学習意欲（子どもたちは学習に能動的に取り組もうとしているか）についてもあわせて検討していく必要がある。

「学力低下」を決定づけたとも言われる 2003 年の PISA 調査の結果を、上記の 4 つの視点を念頭に置いて読み解いてみよう。なお、2015 年調査はコンピュータ使用型調査への移行に伴い、尺度化・得点化の方法の

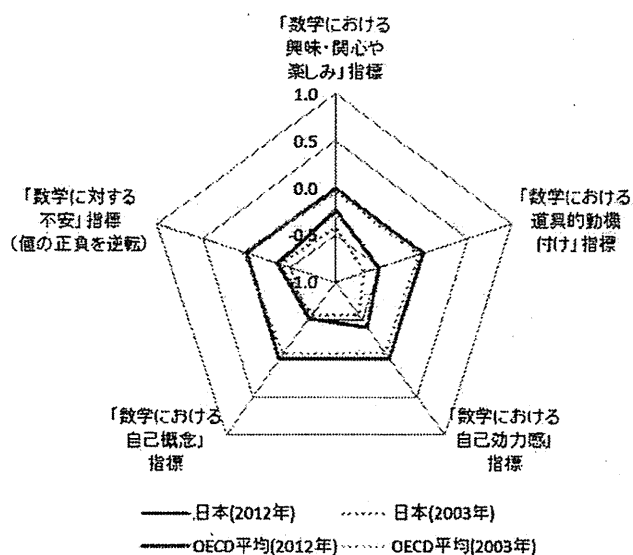
変更等があったため、経年変化については、2012 年調査までの結果に基づく資料により考察する。

資料 1、2 は、PISA 調査の平均得点と順位、および習熟度レベル別割合それぞれの経年変化である¹³。2000 年調査と 2003 年調査を比べてみると、まず日本の学力水準について、資料 1 のように、読解力では、得点と順位の低下がみられる。数学的リテラシーも得点の低下が見られるが、順位としては依然として 1 位グループであった。数学と科学については、2006 年調査では低下傾向が顕著となるものの、2003 年調査の時点では、世界的に高い水準にあったと言える。

次に、学力構造という視点から考える際に、もう一つの国際学力調査である IEA の TIMSS 調査（小学 4 年生と中学 2 年生を対象に、算数・数学と理科の学習到達度を問う）において、日本の子どもたちの学力水準が一貫して高いという事実は示唆的である¹⁴。そこからは、知識・技能の習得を問う従来型の学力ではなく、まさに PISA 型の学力に課題があったということが見えてくる。

これまで日本で読解問題という場合、指示語の内容を答えさせたり、テキストの主題について説明させたりと、ひとつのテキストを細かく読み取っていく問題であった。これに対して、PISA の読解問題では、複数のテキストを読み比べて、自分の意見を述べるのが求められている。また、数学の問題では、メイリンという学生が留学する場面が設定され、通貨の両替に関する計算をしたり、為替レートの変動に伴う損得を判

資料3. 数学的リテラシー得点に影響を与える要因の経年変化



断したりすることを求める問題や、グラフの一部分を示して盗難事件が激増したとするTVレポーターの解説の不適切さを、示されたデータの相対的な意味を指摘して、根拠を明らかにしながら説明する問題が出題されている。PISAでは、知識を習得しているだけでなく、自分の考えを筋道立てて論述したり、知識を実社会・実生活で活用したりする力が問われたのである。

なお、2006年度調査以降、読解力の成績において伸びが顕著なものも、結果としてPISA（その背景にある教育内容のヨーロッパ標準）と日本の教育課程との間で学力観や教科観の違いが調整されたからであろう。PISAショック以降、PISA型学力を意識した取り組みが各教科で行われてきたが、算数・数学にしても理科にしても、それは記述式問題の弱さ（白紙答案の多さ）への対応として、PISA型読解力の教科横断的な育成として矮小化され、「言語活動の充実」という授業レベルでの工夫に止まった。一方で、実社会・実生活をよりよく生きることと結びつけて各教科の内容や系統を見直すこと、すなわち、カリキュラムレベルでの見直しは十分に行われなかった。これに対して、読解指導については、従来の日本の指導方法との違いが目に見えて大きく、しかも、読解については、「いかに学ぶか」（教室での言語活動の形式）が「何を学ぶか」（言語運用能力）と密接に結びついているために、結果として

カリキュラムレベルでの調整（チューニング）がなされた。読解力の成績のV字回復の背景にはそうした要因もあると考えられる。

学力格差という点に関して、資料2は、読解力において、日本はレベル5以上の成績上位層の割合が多い一方で、2000年から2003年にかけて、レベル1未満の下位層の割合が増加したことを示している。こうした下位層が生まれたことも、読解の順位の低下に関係していると推測できよう。学力の水準やその向上を目指す一方で、そうした下位層の底上げという課題を見落としてはならない。しかも、そうした学力格差の傾向は、TIMSSよりもPISAにおいて見られ¹⁵、そこからは、他ならぬ知識を活用する力における学力格差の存在に注意を向ける必要性が見えてくる。

最後に学習意欲に関連して、資料3が示すように、日本は、数学に対して肯定的なイメージや動機を示している生徒の割合が低い。2012年調査では肯定的な回答の割合が上昇しているが、依然としてOECD平均からすると意欲面は低い。こうした日本の子どもたちの意欲面の低さは、さまざまな調査で一貫して指摘されてきた傾向であり、「日本は学力トップレベル、学習意欲はワーストレベル」という状況は、「病める学力」や「日本型高学力」といった形で問題視され続けてきた¹⁶。日本の謙遜の文化の影響など、学習意欲の高低については、文化差を考慮に入れねばならず、他国と機械的に比較することには慎重であるべきだろう。しかし、学力向上を急ぐあまり、「学びからの逃走」を加速させてしまうことには注意が必要である。

こうして、教師や保護者や市民が、自分たちの目でポイントを押さえて学力調査の結果を眺め、簡易な分析により概要をつかむことができれば、「学力低下⇒学力向上」といった、単純で一般大衆にわかりやすい構図を相対化し、国の教育政策の妥当性を検討したり、見落とされがちな問題に着目したりすることが促されるだろう¹⁷。

上記の学力調査を読み解く4つの視点に加えて、学力調査の結果に影響を及ぼす背景要因の問題も考

慮しておくことが必要である。そもそも学力調査が明らかにする「学力水準」については、学習者個人の能力や教師個々人の力量のみに規定されるものではなく、教室の関係性や文化、学校の組織としての力や地域の社会環境などの影響も大きい。「学力水準」は、いわば学校をめぐる社会関係資本の量的・質的豊かさの指標という側面も持っている¹⁸。ゆえに、学力テストの結果の良し悪しを、学校や教師の教育活動の質や努力のみに帰責することには慎重でなければならないし、学力形成を真に実現するには、学習や学校組織の共同性を大切にしなければならないのである。

（３）実践的判断におけるエビデンスの位置

学校教育について語る上で学力調査の存在感が増してきたことは、エビデンスに基づく教育という考え方の浸透と密接に結びついている。学力調査の結果は、教育における有力な実証的エビデンスとして重視されているのである。ゆえに、学力調査の結果を教育実践や教育政策のデザインや改善に生かしていく際には、そもそも実証的エビデンスが教育上の様々な意思決定においてどのような意義と限界を持つのかを認識しておく必要がある¹⁹。

専門職の仕事において、臨床的・実践的判断が重要であることは、エビデンスに基づく実践・政策の推進者も批判者も合意するところである。エビデンスに基づく実践・政策を先導してきた「エビデンスに基づく医療（evidenced-based medicine）」においても、改めて場の状況や患者の好みの重要性が再確認されているということをふまえれば、実践的判断の状況依存的な性格、および、そうした個別具体的な判断の的確さを支える経験に基づく実践知の役割の大きさを否定することはできないだろう。特に教育という仕事については、何がよい教育かという教育目的・目標自体が論争的であり、しかも教えるプロセスは、教師と子ども、子どもたち相互の複雑で象徴的なコミュニケーションによって成立する不確実性の高い営みである。そうした教育という営みの性格をふまえるなら、教職の専門職性の根拠は、その実践が厳密な基礎科学や応用技術に基づいていること以上に、実践の複雑性やそこに働く高度で総合的な見識・判断に、第一義的には求められ

るべきであろう。

ゆえに、これをやればこういう効果が上がるという、一般化された技術知を過度に信頼することは、専門職の実践的判断を阻害する危険性がある。たとえば、メタ分析の結果得られた効果量が低くても、その方法を実施すべきではないということを意味しないし、メタ分析の結果得られた知見は過去に関する見解であり、決定的なものではない²⁰。しかし、さまざまな教育上の決定をする際に、技術知に限らず量的データやそれに裏付けられた科学的根拠（実証的エビデンス）を、その限界を自覚しながら参照することは、実践改善の示唆を与えうる。そしてその有効性の程度は、実践レベルと政策レベルとは異なってくる。

実践レベル、特に教室での授業レベルでは、一瞬一瞬の状況判断や子どもへの応答が求められる。応答の関係性の下で個々人に対する具体的で総合的な判断（熟慮）が求められる実践の場において、もっとも依拠すべきは、実践の中の理論（直観的なものだが、なぜそれが妥当かの事後の説明は可能なもの）である。実証された技術知は、あくまでも実践上のレパートリーを提供するのみである。また、ペーパーテストなどによって得られた量的データの系統的な収集・分析以上に、実際の子どもの表情・行動やノートの記述などから学習の質を多面的に具体的に解釈することが意味を持つ。量的データを活用するにしても、教師自身が、自分がベストと思うタイミングで、ほどほどの厳密性でデータを収集し、簡易な分析を行うことが重視されるべきだろう。

これに対して、教育課程（教育内容の選択と組織化）の枠組みの設定や、教育条件（教具・メディアの整備や学級・学校組織の編成など）の決定に関わる政策レベル、あるいは、学期や年間の学校の取り組みの検証や、特定の教材の有効性を検証したりする、政策と授業の狭間のレベルについては、実証的エビデンスは一定有効である。そうしたレベルの決定は、教えるプロセスを規定する、内容や時間や空間や人の配置などの構造を対象とするものである。そして、統計的手法による大規模調査が明らかにする、集団の全体的な傾向に関する知見は、そうしたマクロな構造の決定において確かな根拠を提供しうる。ただしその際も、たとえば、開発途上国を対象になされた大規模調査の結果

を、日本における少人数学級の費用対効果の問題にそのまま適用するといったことにならぬよう、外的妥当性の問題を考慮する必要がある。

また、政策レベルの決定においては、さまざまな利害関係者の参加による委員会や審議会での熟議が必要であり、量的データのみに依拠するというのではなく、専門家の見解や事例など、弱いとされるエビデンスも含め、多様な根拠に基づいて意志決定がなされることが重要となる。その際、後述するように、エビデンスが適切に生かされる条件として、政策の検証のシステムを確立することと、政策立案のための議論の手順と組織を確立することが同時になされねばならないだろう。たとえば、全国学力テストについても、真に「エビデンスに基づく教育」をめざすのであれば、その結果をどう学習指導要領改訂や条件整備に関する意志決定に生かしていくのかという、教育政策に関する審議のプロセスが同時に問われる必要があるのである。

個々人との相互作用による「実践」を対象にした判断か、教育システム全体の「構造」を対象にした判断かによって、また、意志決定にかけられる時間の長さによって、判断において主に依拠すべき根拠の種類や、量的で実証的な知見の役割の違いが生じる。そして、子どもと向き合う教室レベルに近づくほど、より質的で解釈的な知が、そして、制度の枠組みを決めるマクロな政策レベルに近づくほど、学力調査の結果など、量的で実証的な知が果たす役割が相対的に大きくなるのである。

（４）学力調査の性格規定と設計方針

学力調査の目的と活用に関わっては、PISA 型学力の向上をめざして成果主義的な傾向が強まる中、PISA にしても、全国学力テストにしても、改革や取り組みの成果（事実）を検証する役割以上に、めざす能力像（理念）を提示し改革を先導・牽引する役割を果たしている点に注意が必要である。そもそも PISA は、学校で意図的に指導した学習成果（taught outcome）を調査する「学力」調査では必ずしもなく、学校の学習や学校外での生活に関わらず、生徒が学ばえた能力一般（learned outcome）を調査する「能力」調査として設計されていた。それゆえ、そもそも PISA の結果は、学校カリキュラムの改善だけではなく、学習者の学校

資料４．全国学力・学習状況調査の調査目的

- ・義務教育の機会均等とその水準の維持向上の観点から、全国的な児童生徒の学力や学習状況を把握・分析し、教育施策の成果と課題を検証し、その改善を図る。
- ・そのような取組を通じて、教育に関する組織的な検証改善サイクルを確立する。
- ・学校における児童生徒への教育指導の充実や学習状況の改善等に役立てる。

内外での学習環境全体の再検討とデザインのための手がかりとすべきものなのである。

PISA の理念提示型調査としての性格は、日本の学力論議において、PISA が大きなインパクトを持ち得た理由とも関係している。2000 年前後の学力低下論争では、学力低下の事実（テストスコア）が主な争点となり、学力の中身に関する議論は欠落しがちであった。他方、教育研究者の学力論は理念論のみに陥りがちで、それを具体化したり検証したりする道筋は十分に示されてはいなかった。目標（価値）論なき実証志向の評価論でもなく、評価論なき理念論としての学力論でもなく、教育目標論と教育評価論とを一体のものとして提起した点に、PISA がインパクトを持ち得た理由があると考えられる。

そして近年 PISA は、非認知的要素を含むキー・コンピテンシーそのものの評価をめざして、より包括的な能力を対象化すべく、能力調査としての側面を強めている。より直接的には、それぞれのリテラシーの定義において、学習者の主体的関与（engagement）といった動機付的な側面が強調されるとともに、創造的でインタラクティブな能力を測定すべく、筆記型から ICT を活用した調査への移行も追求されている点に、こうした傾向を見て取ることができる²¹。

また、2012 年調査から数学的プロセスが、能力の階層性を示す「能力クラスター」（再現、関連付け、熟考）から、能力の要素や局面を示す「定式化」「適用」「解釈」へと変更されたことは、PISA が能力調査としての側面を強めていることを象徴的に示しているように思われる²²。平行四辺形の公式は覚えていても、その意味までわかって、斜辺と高さを間違えずに必要な情報を取り出して求積できるとは限らないといった具合に、学力調査として内容に即した学びの深さを診断するには、階層性（能力の質的レベル）に着目する必要がある。それゆえ、TIMSS では、「知ること」「応用するこ

と」「推論を行うこと」といった具合に、認知過程のカテゴリは階層性を示すもので構成されている。他方、PISA は、「熟考」という能力の質的レベルの重要性を提起するものであり、内容習得の層を評価するカテゴリを組み込まず、熟考レベルの思考だからこそ重視されるプロセスの局面（定式化、適用、解釈）のみで評価の枠組みを構成しても、不都合は生じない。むしろ、そうしためざす能力の層のみをより直接的に打ち出す方が、理念提示型調査としてはより効果的に機能するだろう。

だが、学力調査が、改革を先導・牽引する理念提示型調査として機能している状況は、下記のような点において限界がある。まず、大規模調査の制約ゆえに、PISA にしても全国学力テストにしてもペーパーテストがベースであり、ICT などの技術により、相互作用的な調査が可能になったとしても、めざす能力像のすべては示し得ない。実際、次期学習指導要領改訂の議論においては、PISA（認知的要素中心）は現代社会で求められるキー・コンピテンシー（他者対話し協働する力や自律的に学び続けていく力などの非認知的要素も含む）の一部を評価しているに過ぎないとして、非認知的要素も含む汎用的スキルなどの資質・能力を重視する改革が進行している。

また、理念提示型調査は、理念を具体化する教育課程や教育条件の設計と結びつかない場合、調査主体は理念を提示するのみで、その実現の責任を現場に負わせることになりかねない。全国学力テストや教育課程実施状況調査など、国が行う学力調査については、教育現場が主体となって行うべき指導改善という目的に解消されない、カリキュラム評価・政策評価という、大規模調査だからこそ可能な目的が第一義的に追求される必要がある。だが、現行の全国学力テストについては、資料4のように、これら多様な目的が混在しており、第一に考慮されるべきカリキュラム評価・政策評価のための調査として設計されているとはいえない。それどころか、指導改善に生かすことや現場レベルでのPDCAサイクルの強調により、行政レベルでの責任を問うことには結びつかず、現場に主体的な自助努力を強いる傾向にある。教室や学校レベルで学習者個々人の学習状況の把握をめざす心理「検査」論や教育評価論と、制度レベルで社会集団の傾向や国全体の

実態の把握をめざす社会「調査」論や政策評価論とは、同じく「測定」や「評価」という言葉で語られるが、区別して議論すべきものであろう。

第一義的な目的を空洞化させたままの全国規模の学力調査の実施は、実施した結果に引きずられる形で、目的が逆規定される。学力テストの結果を内申点と結びつけることで順位を上げた大阪府の例は、都道府県間の差が縮まっていることも手伝って、都道府県間の「学力向上」競争を激化させ、テスト結果を上げることの自己目的化と、実質的な教育改善や学力保障という理念の空洞化が危惧される。さらに言えば、政治的アピールの道具として学力テストが用いられる傾向が強まることに注意が必要である。

なお、そもそも教育活動や学力形成の結果は、水準が上がったか下がったかで議論されるべきものなのか。人権事項としてすべての子どもたちに一定の内容や能力が保障されたかどうかという形で学力が議論される必要があるのではないか。2000年前後の学力低下論争以降、「学力向上」にしても「学力格差」にしても、学力テストのスコアをもとに、相対的な量的差異や程度の問題として学力問題が語られる（学力向上のレトリック）一方で、子どもの絶対的貧困の問題とも連動した、内容の未習得や低学力（落ちこぼし）に対して、すべての子どもたちに保障されるべき絶対的・普遍的な内容・基準、つまり学力や教育課程の中身を問う視点（学力保障のレトリック）は後景に退くことになった。内容（実体）ではなく、能力（機能）から教育課程を編成しようとする動きは、そうした傾向を強めかねない。

カリキュラム評価・政策評価としての学力調査であれば、必ずしも悉皆調査である必要はなく、悉皆調査で行うにしても、採点や子どもたちへのフィードバックは各学校にゆだね、国は一部を抽出して分析を行う形も考えられるだろう。さらに言えば、カリキュラム評価・政策評価という目的を意識するのであれば、先述のように、学習指導要領改訂のプロセスや政策決定過程のデザインと一体のものとして、学力調査のあり方が議論される必要があるだろう。学力調査は是か非かという二項対立の議論を超えて、今こそ、学力調査を教育政策・教育実践の改善と学力保障の道具として設計していくことが求められる。

※本稿は、拙稿「学力調査の時代を読み解く評価リテラシー」『科学研究費補助金・研究成果最終報告書・思考力・判断力・表現力育成のための長期的ルーブリックの開発』2016年3月に若干の修正を加えたものである。

注

¹ 犬山市教育委員会編『全国学力テスト、参加しません。』明石書店、2007年。

² 北野秋男・下司晶・小笠原喜康『現代学力テスト批判』東信堂、2018年などを参照。

³ 佐貫浩・世取山洋介編『新自由主義教育改革：その理論・実態と対抗軸』大月書店、2008年などを参照。

⁴ 石井英真『増補版・現代アメリカにおける学力形成論の展開—スタンダードに基づくカリキュラムの設計—』東信堂、2015年の補論などを参照。

⁵ OECD（渡辺良監訳）『PISA から見る、できる国・頑張る国—トップを目指す教育』明石書店、2011年、福田誠治『ネオリベラル期教育の思想と構造』東信堂、2017年などを参照。

⁶ 田中耕治編『グローバル化時代の教育評価改革—日本・アジア・欧米を結ぶ』日本標準、2016年を参照。

⁷ 志水宏吉・高田一宏編著『学力政策の比較社会学・国内編：全国学力テストは都道府県に何をもたらしたか』明石書店、2012年。

⁸ 石井英真「学力向上」篠原清昭編著『学校改善マネジメント』ミネルヴァ書房、2012年を参照。

⁹ 石井英真「現代日本の学力向上政策の検討—『スタンダードに基づく教育改革』の日本の特質」『日本デュイ学会紀要』第54号、2013年を参照。

¹⁰ 川口俊明「学力調査をととした『統制』を論じるだけでなく」『教育』2017年11月でも、学力調査としての適切性自体を検討することの必要性が述べられている。

¹¹ 木村拓也「戦後日本において『テスト専門家』とは一体誰であったのか？」『教育情報学研究』第4号、2006年。

¹² テスト理論や測定論の観点からの評価リテラシーの実実を考える上では、日本テスト学会編『テスト・スタンダード—日本のテストの将来に向けて』金子書房、2007年、日本テスト学会編『見直そう、テストを支える基本の技術と教育』金子書房、2010年、光永悠彦『テストは何を測るのか』ナカニシヤ出版、2017年などが参考になる。

¹³ 資料1は、国立教育政策研究所『OECD 生徒の学習到達度調査（PISA2012）のポイント』2013年

（http://www.nier.go.jp/kokusai/pisa/pdf/pisa2012_result_point.pdf：2019年3月13日確認）より、資料2は、同『OECD 生徒の学習到達度調査—2012年調査国際結果の要約』2013年

（http://www.nier.go.jp/kokusai/pisa/pdf/pisa2012_result_outline.pdf：2019年3月13日確認）より、資料3は、同『OECD 生徒の学習到達度調査—2012年調査分析資料集』2013年（http://www.nier.go.jp/kokusai/pisa/pdf/pisa2012_reference_material.pdf：2019年3月13日確認）より抜粋。

¹⁴ 田中耕治編『新しい学力テストを読み解く—PISA/TIMSS/全国学力・学習状況調査/教育課程実施状況調査の分析とその課題』日本標準、2008年を参照。

¹⁵ 同上書を参照。

¹⁶ 駒林邦男『改訂版・現代社会の学力』放送大学教育振興会、1999年、佐藤学『学びから逃走する子どもたち』岩波書店、2000年などを参照。

¹⁷ 川口俊明「国際学力調査からみる日本の学力の変化」『福岡教育大学紀要』第63号、第4分冊、2014年では、PISAの結果について、調査・統計の専門的な知見をふまえた厳密な検討がなされている。たとえば、読解力の回復に関する解釈も、PISAの調査設計に伴う問題であった点が指摘されている。また、家庭環境に関する調査結果も用いながら、格差の拡大についても検討し、中学校の数学と理科で拡大傾向を指摘している。

¹⁸ 志水宏吉『「つながり格差」が学力格差を生む』亜紀書房、2014年などを参照。

¹⁹ 石井英真「教育実践の論理から『エビデンスに基づく教育』を問い直す—教育の標準化・市場化の中で—」『教育学研究』第82巻第2号、2015年などを参照。

²⁰ 山森光陽「監訳者解説」ジョン・ハッティ（山森光陽・監訳）『教育の効果』図書文化、2018年。

²¹ 国立教育政策研究所編『生きるための知識と技能 5—OECD 生徒の学習到達度調査（PISA）2012年調査国際結果報告書』明石書店、2013年を参照。

²² 階層性としての能力概念と要素としての能力概念については、石井英真『今求められる学力と学びとは』日本標準、2015年を参照。

（教育方法学研究室 准教授）

受理 2019年3月12日